МІНІСТЕРСТВО ОХОРОНИ ЗДОРОВ'Я
НАЦІОНАЛЬНИЙ МЕДИЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ О. О. БОГОМОЛЬЦЯ

ФАРМАЦЕВТИЧНИЙ ФАКУЛЬТЕТ
КАФЕДРА АНАЛІТИЧНОЇ, ФІЗИЧНОЇ ТА КОЛОЇДНОЇ ХІМІЇ

ВИПУСКНА МАГІСТЕРСЬКА РОБОТА

на тему **«Prediction of toxicity of phenols using artificial neural networks»**

Виконав: здобувач вищої освіти 5 курсу
факультету підготовки іноземних громадян
групи 7602Фа
напряму підготовки 7.12020101 «Фармація»
освітня програма «Фармація»
Mustafa Jolo Ali Al Saker

Керівник: к.х.н., доцент Пушкарьова Ярослава Миколаївна

Рецензент: к.х.н., доцент, доцент кафедри хімічного матеріалознавства хімічного факультету Харківського національного університету імені В.Н. Каразіна
А.В. Пантелеймонов

КИЇВ-2022

CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

ABSQon − sum of absolute charges on nitrogen and oxygen atoms in a molecule

ANN – artificial neural network

E_{LUMO} − the energy of the lowest unoccupied molecular orbital

IGC_{50} − the concentration in mmol/L of the toxicant causing 50% inhibition of growth to Tetrahymena pyriformis

Log *D* – distribution coefficient

MaxHp – the largest positive charge on a hydrogen atom

MW – molecular weight

P − the 1-octanol/water partition coefficient

P_{NEG} − is the negatively charged molecular surface area in percent's

pK_a – negative logarithm of the ionization constant

PLS – regression on partial least squares

RBFN − radial basis function network with addition of one neuron at a time

RBEFN − radial basis function network with zero error on training vectors

QSAR − quantitative structure-activity relationship

SsOH – electrotopological state index for the hydroxyl group

INTRODUCTION

Actuality. Determining the toxicity of chemicals is one of the most important stages along the way creation of medicines. This the indicator is of great importance not only in pharmacology, but also in industry and many other areas of human activity where there is potential contact with harmful substances − agriculture, perfumes, detergents, etc. It is known that the experimental study of only one type of toxicity requires a large number of animals, considerable time and is time consuming. This was facilitated by the high cost of experimental studies in toxicology. In studies of various aspects of experimental determination of toxicity, it becomes very important to use calculation methods to predict these indicators, which allows to assess in advance the possible risk of using chemicals without additional experiments.

Phenolic compounds are characterized by different medical and health uses (antioxidant effect, antibacterial effect, anti-cancer effect, cardioprotective effects). That is why they are interesting from a toxilogical point of view. There are different mechanisms of toxic action of phenols namely polar narcotics, weak acid respiratory uncouplers, pro-electrophiles and soft-electrophiles.

Artificial neural network (ANN) is a flexible mathematical model. There exist numerous applications of ANNs in data analysis, pattern recognition, adaptive control, prediction, classification, identification, etc.

Given the above we can include, that using artificial neural networks for prediction the toxicity of phenols is actual and useful topic.

The aim of the work is search of optimal parameters of artificial neural networks to ensure high reliability of predicting the toxicity of phenols, which will contribute to the further development of computational methods for predicting the toxicity of chemical compounds.

Achieving this goal determines the solution of the **following tasks**:

1) choose the optimal number of hidden neurons for construction of feed-forward neural network and cascade neural network;

2) choose the optimal spread value of the Gaussian transfer function for radial basis function neural networks;

3) estimate the effectiveness of the application of artificial neural networks algorithms to predict the toxicity of chemical compounds by using mean squared error;

4) compare the obtained results with the results found in literature (results of using quantitative structure-activity relationships (QSARs) and regression on partial least squares algorithms (PLSs)).

Subject of study is parameters of algorithms of artificial neural networks.

Object of study is toxicity of phenols based on data on seven molecular descriptors.

Research methods are algorithms of artificial neural networks.

Practical significance of the obtained results. Obtained results can be useful for development computational methods, which allows to assess in advance the possible risk of using chemicals without additional experiments. The results of this work were presented in the XII International Scientific and Practical Conference “Actual priorities of modern science, education and practice”, March 29-April 01, 2022, Paris, France and III International Scientific and Theoretical Conference “Theory and practice of modern science”, April 1, 2022, Krakow, Republic of Poland (certificates of participation are presented as supplementary information).

Scientific novelty. Artificial neural network algorithms (feed-forward neural network, cascade neural network and radial basis function neural networks) were first used to predict the toxicity of phenols based on a set of molecular descriptors. It is shown that feed-forward neural network is the most effective algorithm among used ANN algorithms for prediction the toxicity of phenols. It is shown that the results obtained by using of feed-forward neural network are characterized by higher accuracy than results obtained by using of QSARs and PLSs.

SECTION 1. LITERATURE REVIEW

1.1 Methods for the toxicity prediction and evaluation of phenols

Determining the toxicity of chemicals is one of the most important stages along the way creation of medicines. This the indicator is of great importance not only in pharmacology, but also in industry and many other areas of human activity where there is potential contact with harmful substances − agriculture, perfumes, detergents, etc. It is known that the experimental study of only one type of toxicity requires a large number of animals, considerable time and is time consuming. Computer prediction of the toxicity of chemical compounds began to develop in the 1980s. This was facilitated by the high cost of experimental studies in toxicology. In studies of various aspects of experimental determination of toxicity, it becomes very important to use calculation methods to predict these indicators, which allows you to assess in advance the possible risk of using chemicals without additional experiments [1].

Phenolic compounds are characterized by different medical and health uses (antioxidant effect, antibacterial effect, anti-cancer effect, cardioprotective effects). That is why they are interesting from a toxilogical point of view. There are different mechanisms of toxic action of phenols namely polar narcotics, weak acid respiratory uncouplers, pro-electrophiles and soft-electrophiles. By far the largest number of toxicity data are available for the inhibition of growth to the protozoan ciliate Tetrahymena pyriformis [2−5].

Different approaches for classification and prediction of the toxicity of phenols was used. There have been many attempts to develop QSARs (quantitative structure activity relationship) for the prediction of the toxicity of phenolic compounds.

In 1996 Cronin and Schultz [6] were able to develop a two-parameter QSAR for the prediction of phenols toxicity to Tetrahymena pyriformis based on descriptors for hydrophobicity and electrophilicity:

$$\log 1/IGC_{50} = 0.671(\pm 0.022)logP - 0.670(\pm 0.055)E_{LUMO} - 1.123, \quad (1.1)$$

$$n = 120,\ r^2 = 0.899,\ r_{CV}^2 = 0.893,\ s = 0.262,\ F = 523,$$

where IGC_{50} is the concentration in mmol/L of the toxicant causing 50% inhibition of growth to Tetrahymena pyriformis,

P is the octanol–water partition coefficient,

E_{LUMO} is the energy of the lowest unoccupied molecular orbital,

n is the number of observations,

r^2 is the correlation coefficient,

r_{CV}^2 is the cross-validated correlation coefficient using a leave one-one-out approach,

s is the standard error of the estimate,

F is the Fisher criterion and figures in parentheses are the standard errors on the coefficients.

Garg et al. [7] in 2001 demonstrated a similar relationship to equation 1, but replaced LUMO with Hammett constant σ:

$$\log 1/IGC_{50} = 0.64(\pm 0.04)logP + 0.61(\pm 0.12)\sigma + 1.84\ (\pm 0.13), \quad (1.2)$$

$$n = 119,\ r^2 = 0.896,\ r_{CV}^2 = 0.887,\ s = 0.265,\ F \text{ not given.}$$

Mark Cronin et al. in [5] have proposed and compared different approaches for developing of QSARs, which are used for the prediction of 200 phenols toxicity to Tetrahymena pyriformis. Among them are response-surface approach or two-parameter approach, stepwise regression or seven-parameter approach, two dimension and three dimension partial least squares. Disadvantages of presented in [5] approaches are observation of outliers. But in general all these approaches was found to be a good model for solving the task of prediction of the phenols toxicity.

Aynur Aptula et al. [8] have used the stepwise linear discriminant analysis (LDA) for classification of the toxic mechanisms of action for 221 phenols to Tetrahymena pyriformis. Using the linear discriminant analysis for classification implies the presence of information about a priori groups of compounds. That is why 221 phenols were a

priori grouped into the four classes according to the different mechanisms of toxic action. This classification was based on molecular descriptors such as hydrophobicity with and without correction for ionisation, acidity constant, frontier orbital energies and hydrogen-bond donor and acceptor counts. Results of using the linear discriminant analysis are 86−89% overall correct classification of phenols into four classes according to the mechanisms of their toxic actions.

Dieguez-Santana et al. [9] were used the multiple linear regression technique to develop a linear quantitative-structure toxicity relationship (QSTR) model for prediction of phenols toxicity to Tetrahymena pyriformis. The obtained model was statistically significant and robust indicating the capability of predicting the aquatic toxicity of phenol derivatives in the impairment of the population growth of *Tetrahymena* pyriformis.

Abbasitabar and Zare-Shahabadi [10] were used genetic algorithm and decision tree-based method for prediction of toxicity of phenols to Tetrahymena pyriformis. The advantage of proposed algorithm is that one can use the resultant tree to predict phenol toxicity with high accuracy with no a priori knowledge about chemical class or mechanism of action of phenols.

Chen et al. [11] were used popular classification algorithm random forest learner for in silico prediction of toxic action mechanisms of phenols to Tetrahymena pyriformis. One global and four local classification models were constructed by employing random forest in the cost-sensitive learning framework. The statistical results in the paper confirmed that random forest was a competitive tool for building classification models of toxicity mechanisms prediction.

Ren [12] was investigated the possibility of using the decision tree-based approach for classification as well as for prediction of toxic action mechanisms of phenols. It was created a three level decision tree with six terminal nodes. This decision tree-based approach achieved prediction accuracy 85%.

1.2 Medical and health uses for phenolic compounds

Phenols form a large and structurally diverse group of compounds. Phenolic compounds are own defined as compounds that possess an aromatic ring with at least one hydroxyl group, and their structure can vary from simple molecule to complex polymer with high molecular weight. These compounds are widely used both in industry and as consumer products (components of dyes, polymers, pharmaceuticals and other organic substances, textiles, leather, paper, oil). They are interesting from a toxilogical point of view. The toxicity of phenols involves a number of different mechanisms of toxic action, respiratory uncouplers and electrophilicity [8, 13, 14].

Phenol is toxic in pure form. Despite this fact phenol have many medicinal applications. Such as injection phenol is used for treatment of muscle spasticity. Phenol is ingredient of vaccine preservative that is used for contaminating the vaccine solutions. Phenolic compounds are characterized by numerous benefits for human health. One of the most important of them is antioxidant property. They are naturally occurring compounds present in many foods, including fruits, vegetables, cereals [2, 3, 15]. Phenolic compounds are divided into phenolic acids, flavonoids, lignans and stilbenes according to the chemical structures [4].

The numerous applications of phenolic compounds especially flavonoids on human health treatments and laboratory using include [2−4, 16]:

− antioxidant effect;

− antibacterial effect (for example, against skin acne problems);

− anti-cancer effect (numerous studies validated that polyphenols are responsible for lowering tumor growth);

− cardioprotective effects (it was investigated that the consumption of polyphenols minimizes the risk of coronary heart diseases);

− immune system promoting and anti-inflammatory effects (phenolic compounds have demonstrated anti-inflammatory properties to treat skin diseases, rheumatoid arthritis, and inflammatory bowel disease);

− skin protective effect from UV radiation;

− antidiabetic effect.

The use of phenolic compounds are the promising candidate for future medical and pharmaceutical product development as an ingredient to promote human health.

1.3 Artificial neural networks: basics

Artificial neural network algorithms are modern mathematical models that are simulates human brain functioning.

Comparison between a biological neuron and an artificial neural network neuron is shown in Figure 1.1 [18, p. 205].

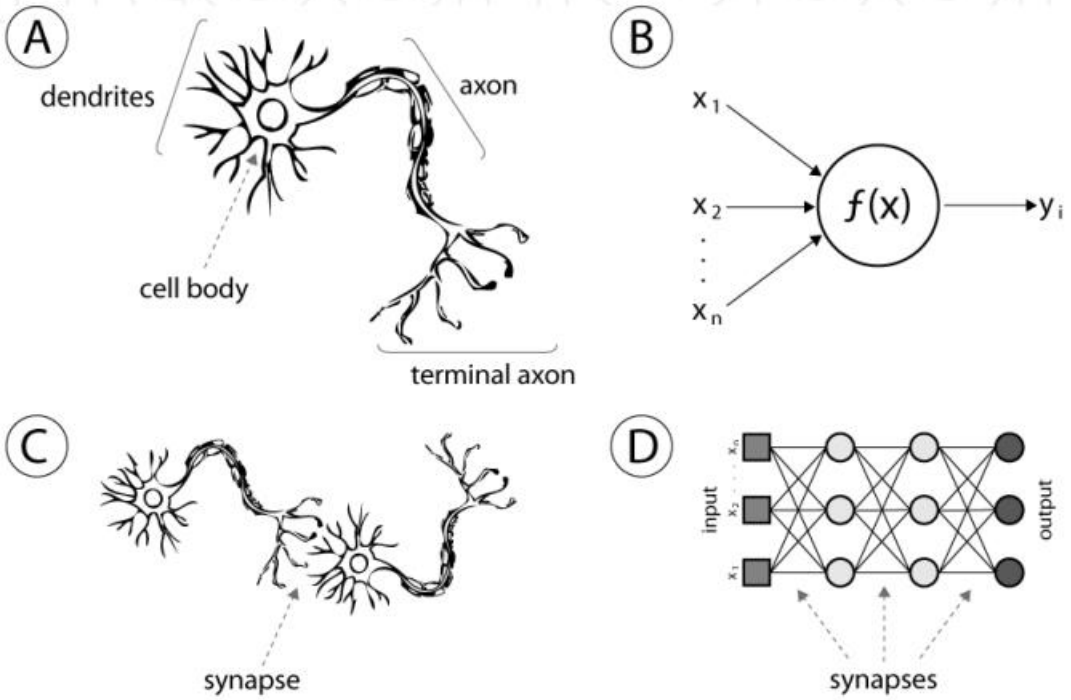

Figure 1.1. Comparison between a biological neuron and an artificial neural network neuron: A) biological neuron; B) artificial neural network neuron; C) biological synapse; D) artificial neural network synapses [18, p. 205]

Many different architectures and types of artificial neural networks are known. Artificial neural networks consist of neurons (artificial neurons). Other name of neurons is hidden units. Artificial neurons are connected with weights and form input, hidden and output layers. Neurons process information and these signals are transfer to the next layer by means of linear or non-linear activation functions [17, 18].

One should choose the transfer or activation functions in each layer, the learning rule, and the number of neurons in each layer for constructing the architecture of an artificial neural network. The input signals multiplied by the weight parameters are summed and passed through a transfer function to produce the output for neurons [19−21].

ANN is a flexible mathematical model. There exist numerous applications of ANNs in data analysis, pattern recognition, adaptive control, etc. [22−24].

Let consider some types of ANNs.

The radial basis function network is among the most commonly used types of ANNs [24−27]. The radial basis function network is a supervised three-layered feed-forward ANN that uses radial basis functions as activation functions. The radial basis function network was formulated initially by Broomhead and Lowe [28].

The Elman network and dynamic neural network are the recurrent networks. The prominent feature of the recurrent network architecture is the presence of feedback or blocks of a dynamic delay. The Elman neural network is a two-layer network with the feedback from the output layer to the input of the hidden layer. Dynamic neural network is the feed-forward input-delay back propagation network [21, 29].

Probabilistic neural network is one of modifications of a radial basis function network. Usually, a spherical Gaussian basis function is used, although many other functions work equally well. Every hidden neuron is intended for single pattern storage of training set. The output layer of probabilistic neural network is the competitive layer which is used to determine the most likely class for a given input vector [21, 30].

Feed-forward neural network is the simplest and widely most useful type of the artificial neural networks. The signals moves from the input neurons, through the hidden neurons, and to the output neurons. This feature provided the name of this type of artificial neural network. In cascade neural network hidden neurons are added to the network once and keep unchanged in the afterwards [21, 31, 32].

The Kohonen network is a simple two-layer unsupervised network. This type of ANN uses a competitive learning algorithm. During the training stage, the input vector is presented to the network and only one neuron (winning neuron) is activated. The

winning neuron is selected as the neuron that has the smallest Euclidean distance to the input vector. Weight coefficients of the winning neuron are modified according to the learning rule. A comprehensive description of the competitive learning algorithm was published elsewhere. The Kohonen network is intended for the classification of input vectors into groups; the number of classes must be assigned *a priori* [33, 34]

Let present some examples of applying ANN techniques for solving pharmaceutical problems: virtual screening for classifying of compound database [35, 36]; control the process of the pharmaceutical production [37−39]; the drug design research [40−42].

Conclusions to section 1

1. In studies of various aspects of experimental determination of toxicity, it becomes very important to use calculation methods to predict these indicators, which allows you to assess in advance the possible risk of using chemicals without additional experiments.

2. Phenolic compounds are widely used both in industry and as consumer products. They are interesting from a toxilogical point of view. The toxicity of phenols involves a number of different mechanisms of toxic action, respiratory uncouplers and electrophilicity.

3. Artificial neural networks have received much more attention recently. Thanks to their adaptive structure and learning capability, they are success fully used to solve classification, identification and prediction tasks.

SECTION 2. EXPERIMENTAL PART

2.1 Data set

Artificial neural networks were developed on a training subset of 197 phenols, and validated by an external testing subset of 20 compounds [5]. The validation subset (testing subset) of compounds was selected prior to ANN model development, and these data were not used in the development of models. All chemicals and their toxicity are listed in Tables 2.1 and Table 2.2.

The toxicity values were from a population growth impairment test using the ubiquitous freshwater ciliate *Tetrahymena pyriformis* (strain GL-C), performed following the protocol previously described by Schultz in 1997 [43]. The 50% growth inhibition concentration, IGC_{50}, was determined for each compound using the Probit Analysis routine in the Statistical Analysis System (SAS) software (SAS Institute 1989). All statistical analyses were performed on nominal concentrations; chemical analyses of concentrations were not performed.

Physico-chemical descriptors which are calculated for the phenols and listed in Tables 2.1, 2.2:

1) Log *D* – distribution coefficient at pH = 7.35, which was calculated according to the expression:

$$\log D = \log P + log(1 + 10^{pH - pK_a}), \qquad (2.1)$$

where log *P* – logarithm of the 1-octanol/water partition coefficient, pK_a – negative logarithm of the ionization constant;

2) E_{LUMO} – energy of the lowest unoccupied molecular orbital;

3) MW – molecular weight;

4) P_{NEG} − is the negatively charged molecular surface area in percent's;

5) ABSQon − sum of absolute charges on nitrogen and oxygen atoms in a molecule;

6) MaxHp – the largest positive charge on a hydrogen atom;

7) SsOH – electrotopological state index for the hydroxyl group.

Table 2.1. Chemicals, their toxicity to *Tetrahymena pyriforms* and significant physico-chemical descriptors (training subset)

N	Compound	Toxicity	Log *D*	E_{LUMO}	MW	P_{NEG}	ABSQon	MaxHp	SsOH
1	4-Hydroxyphenylacetic acid	-1,50	-2,41	0,140	152,16	40,44	1,007	0,229	17,22
2	3-Hydroxybenzyl alcohol	-1,04	0,30	0,388	124,15	47,97	0,784	0,210	17,38
3	4-Carboxylphenol	-1,02	-1,36	-0,482	138,13	39,71	1,099	0,221	17,14
4	3-Hydroxy-4-methoxybenzyl alcohol	-0,99	0,00	0,338	154,18	42,21	1,121	0,219	17,86
5	4-Hydroxy-3-methoxybenzylamine	-0,97	0,01	0,266	153,2	38,43	1,053	0,181	9,16
6	4-Hydroxyphenethyl alcohol	-0,83	0,62	0,333	138,18	35,58	0,754	0,217	17,39
7	3-Carboxylphenol	-0,81	-1,77	-0,574	138,13	38,7	1,117	0,200	17,20
8	4-Hydroxybenzamide	-0,78	0,23	-0,177	137,15	41,47	1,103	0,219	8,79
9	4-Hydroxy-3-methoxybenzyl alcohol	-0,70	0,00	0,412	154,18	44,33	1,116	0,21	17,84
10	2,6-Dimethoxyphenol	-0,60	0,77	0,388	154,18	42,64	1,105	0,177	9,34
11	2,4,6-Tris(dimethylamino-methyl)phenol	-0,52	-0,75	0,425	265,45	32,71	1,438	0,217	10,38
12	Salicylic acid	-0,51	-2,28	-0,59	138,13	36,61	1,117	0,200	17,31
13	2-Methoxyphenol	-0,51	1,19	0,391	124,15	41,6	0,754	0,173	8,99
14	5-Methylresorcinol	-0,39	1,22	0,341	124,15	43,34	0,748	0,219	17,67
15	4-Methylcyanophenol	-0,38	0,71	0,063	133,16	45,98	0,566	0,217	8,84
16	3-Hydroxyacetophenone	-0,38	1,38	-0,459	136,16	35,77	0,816	0,175	8,91
17	2-Ethoxyphenol	-0,36	1,94	0,422	138,18	39,16	0,751	0,173	9,12
18	4-Acetylphenol	-0,30	1,35	-0,38	136,16	38,63	0,793	0,22	8,83
19	3-Ethoxy-4-methoxyphenol	-0,30	1,78	0,318	168,21	41,16	1,097	0,179	9,12
20	2-Methylphenol	-0,29	0,44	0,370	108,15	34,18	0,394	0,166	8,92
21	2-Hydroxybenzamide	-0,24	1,37	-0,265	137,15	41,05	1,125	0,212	8,98
22	Phenol	-0,21	1,48	0,398	94,12	39,97	0,394	0,166	8,63
23	4-Methylphenol	-0,18	1,94	0,431	108,15	36,58	0,394	0,166	8,76
24	4-Hydroxy-3-methoxyphenethyl alcohol	-0,18	0,33	0,324	168,21	36,72	1,188	0,209	17,86
25	3-Acetamidophenol	-0,16	0,73	0,210	151,18	40,99	1,047	0,221	8,97
26	3-Hydroxy-4-methoxybenzaldehyde	-0,14	0,98	-0,489	152,16	42,85	1,148	0,221	9,14

27	4-Hydroxy-3-methoxyaceto-phenone	-0,12	1,32	-0,404	166,19	41,15	1,103	0,175	9,18
28	3,5-Dimethoxyphenol	-0,09	1,42	0,415	154,18	46,71	1,094	0,220	9,10
29	2-Hydroxyethylsalicylate	-0,08	1,52	-0,475	182,19	43,04	1,476	0,209	17,61
30	3-Methylphenol	-0,06	1,94	0,394	108,15	37,11	0,394	0,166	8,81
31	Methyl-3-hydroxybenzoate	-0,05	1,88	-0,485	152,16	45,04	1,105	0,177	8,95
32	3-Methoxy-4-hydroxybenzaldehyde	-0,03	1,05	-0,478	152,16	41,73	1,148	0,221	9,09
33	4-Hydroxy-3-methoxybenzonitrile	-0,03	1,55	-0,429	149,16	48,16	0,954	0,172	9,10
34	3-Ethoxy-4-hydroxybenzaldehyde	0,01	1,61	-0,452	166,19	39,95	1,157	0,181	9,22
35	4-Fluorophenol	0,02	1,77	0,059	112,11	46,18	0,394	0,166	8,59
36	2-Cyanophenol	0,03	1,21	-0,509	119,13	46,77	0,602	0,172	8,89
37	5-Fluoro-2-hydroxyacetophenone	0,04	2,45	-0,786	154,15	40,83	0,738	0,173	9,02
38	2,4-Dimethylphenol	0,07	2,40	0,399	122,18	37,69	0,394	0,166	9,04
39	2-Hydroxyacetophenone	0,08	1,96	-0,517	136,16	36,51	0,748	0,168	9,06
40	2,5-Dimethylphenol	0,08	2,40	0,347	122,18	35,50	0,394	0,166	9,10
41	Methyl-4-hydroxybenzoate	0,08	1,81	-0,397	152,16	41,62	1,086	0,221	8,86
42	3,5-Dimethylphenol	0,11	2,40	0,387	122,18	32,21	0,394	0,166	8,99
43	4^0-Hydroxypropiophenone	0,12	1,91	-0,443	150,19	34,46	0,793	0,22	9,02
44	2,3-Dimethylphenol	0,12	2,40	0,374	122,18	39,02	0,394	0,166	9,10
45	3,4-Dimethylphenol	0,12	2,40	0,436	122,18	37,93	0,394	0,166	8,94
46	2-Ethylphenol	0,16	2,47	0,386	122,18	35,87	0,394	0,166	9,11
47	Syringaldehyde	0,17	0,73	-0,505	182,19	44,45	1,454	0,179	9,44
48	Salicylhydrazide	0,18	0,58	-0,443	152,17	36,12	1,260	0,228	9,10
49	2-Chlorophenol	0,18	2,01	0,030	128,56	39,39	0,393	0,166	8,79
50	4-Hydroxy-2-methylacetophenone	0,19	1,83	-0,290	150,19	37,40	0,748	0,168	9,01
51	4-Ethylphenol	0,20	2,47	0,435	122,18	30,65	0,394	0,166	8,85
52	3-Ethylphenol	0,23	2,47	0,402	122,18	33,36	0,394	0,166	8,94
53	Salicylaldoxime	0,25	1,87	-0,312	137,15	44,58	0,903	0,208	17,14
54	2,3,6-Trimethylphenol	0,28	2,86	0,382	136,21	35,85	0,359	0,217	9,39

55	2,4,6-Trimethylphenol	0,28	2,86	0,431	136,21	33,26	0,359	0,217	9,33
56	2-Hydroxy-5-methylacetophenone	0,31	2,42	-0,483	150,19	38,52	0,747	0,168	9,18
57	2-Bromophenol	0,33	2,64	-0,049	173,01	36,02	0,394	0,166	8,87
58	5-Bromo-2-hydroxybenzylalcohol	0,34	1,31	-0,007	203,04	33,63	0,786	0,210	17,72
59	2,3,5-Trimethylphenol	0,36	2,86	0,358	136,21	33,03	0,360	0,217	9,28
60	3-Methoxysalicylaldehyde	0,38	1,34	-0,454	152,16	41,23	1,163	0,18	9,23
61	Salicylhydroxamic acid	0,38	0,47	-0,584	153,15	39,71	1,255	0,245	17,24
62	2-Chloro-5-methylphenol	0,39	2,48	0,019	142,59	35,97	0,355	0,218	8,97
63	4-Allyl-2-methoxyphenol	0,42	2,20	0,393	164,22	36,76	0,731	0,219	9,25
64	2-Hydroxybenzaldehyde	0,42	1,55	-0,433	122,13	39,53	0,819	0,175	8,88
65	2,6-Difluorophenol	0,47	1,69	-0,321	130,1	45,75	0,379	0,175	8,46
66	Ethyl-3-hydroxybenzoate	0,48	2,41	-0,453	166,19	42,38	0,973	0,181	9,02
67	4-Cyanophenol	0,52	1,47	-0,413	119,13	47,37	0,602	0,172	8,74
68	4-Propyloxyphenol	0,52	2,37	0,330	152,21	36,45	0,732	0,219	18,27
69	4-Chlorophenol	0,55	2,43	0,095	128,56	33,48	0,394	0,166	8,70
70	Ethyl-4-hydroxybenzoate	0,57	2,35	-0,367	166,19	39,73	1,083	0,221	8,92
71	5-Methyl-2-nitrophenol	0,59	1,83	-1,153	153,15	31,03	0,359	0,217	9,10
72	2-Bromo-4-methylphenol	0,60	2,91	-0,012	187,04	32,46	0,392	0,167	9,00
73	2,4-Difluorophenol	0,60	1,98	-0,318	130,1	40,10	0,379	0,176	8,50
74	3-Isopropylphenol	0,61	2,82	0,415	136,21	31,28	0,394	0,166	9,06
75	5-Bromovanillin	0,62	1,39	-0,702	231,05	41,27	1,163	0,180	9,33
76	*a*; *a*; *a*-Trifluoro-4-cresol	0,62	2,46	-0,348	162,12	39,49	0,394	0,166	8,66
77	Methyl-4-methoxysalicylate	0,62	2,43	-0,428	182,19	45,92	1,424	0,179	9,34
78	4-Bromophenol	0,68	2,49	0,020	173,01	34,76	0,394	0,166	8,74
79	2-Chloro-4,5-dimethylphenol	0,69	2,95	0,053	156,62	35,14	0,384	0,173	9,09
80	4-Butoxyphenol	0,70	2,90	0,330	166,24	32,98	0,732	0,219	8,97
81	4-Chloro-2-methylphenol	0,70	2,89	0,080	142,59	30,52	0,394	0,166	8,99
82	3-*Tert*-butylphenol	0,73	3,17	0,431	150,24	29,49	0,394	0,166	9,18

83	2,6-Dichlorophenol	0,73	2,11	-0,259	163,00	31,36	0,388	0,169	8,94
84	2-Methoxy-4-propenylphenol	0,75	3,00	-0,041	164,22	37,8	0,734	0,219	9,25
85	3-Chloro-5-methoxyphenol	0,76	2,64	0,027	158,59	38,27	0,749	0,175	8,96
86	4-Chloro-3-methylphenol	0,80	2,89	0,133	142,59	35,76	0,394	0,166	8,88
87	2-Isopropylphenol	0,80	2,82	0,408	136,21	33,08	0,394	0,166	9,28
88	2,6-Dichloro-4-fluorophenol	0,80	1,53	-0,568	180,99	25,83	0,380	0,175	8,90
89	4-Iodophenol	0,85	2,91	0,024	220,01	34,63	0,394	0,166	8,75
90	2,2^{0}-Biphenol	0,88	1,48	-0,239	186,22	42,72	0,788	0,166	8,63
91	4-*Tert*-butylphenol	0,91	3,17	0,471	150,24	30,7	0,360	0,217	9,02
92	3,4,5-Trimethylphenol	0,93	2,86	0,430	136,21	37,65	0,360	0,217	9,12
93	2,2^{0},4,4^{0}-Tetrahydroxybenzophenone	0,96	2,64	-0,786	246,23	45,66	1,923	0,221	8,85
94	4-*Sec*-butylphenol	0,98	3,35	0,445	150,24	29,41	0,360	0,217	9,01
95	3-Hydroxydiphenylamine	1,01	2,62	0,104	185,24	42,67	0,610	0,216	9,02
96	4-Hydroxybenzophenone	1,02	2,81	-0,485	198,23	40,89	0,744	0,167	9,10
97	2,4-Dichlorophenol	1,04	2,91	-0,245	163,00	26,87	0,390	0,169	8,85
98	2,4,6-Tribromoresorcinol	1,06	2,74	-0,61	346,79	32,56	0,757	0,218	18,47
99	Benzyl-4-hydroxyphenyl ketone	1,07	2,44	-0,375	212,26	40,24	0,750	0,166	18,28
100	4-Chloro-3-ethylphenol	1,08	3,42	0,141	156,62	29,57	0,390	0,170	9,01
101	2-Phenylphenol	1,09	2,94	-0,119	170,22	40,93	0,359	0,217	9,56
102	2,5-Dichlorophenol	1,13	2,66	-0,325	163,00	24,11	0,389	0,169	8,88
103	3-Chloro-4-fluorophenol	1,13	2,59	-0,264	146,55	38,47	0,383	0,174	8,69
104	3-Bromophenol	1,15	2,62	-0,074	173,01	32,25	0,394	0,166	8,78
105	6-*Tert*-butyl-2,4-dimethylphenol	1,16	4,09	0,455	178,3	27,88	0,359	0,217	9,86
106	4-Chloro-3,5-dimethylphenol	1,20	3,35	0,147	156,62	35,03	0,394	0,166	9,06
107	2-Hydroxybenzophenone	1,23	3,39	-0,629	198,23	41,96	0,810	0,175	19,06
108	4-*Tert*-pentylphenol	1,23	3,70	0,470	164,27	27,15	0,360	0,217	9,10
109	4-Bromo-3,5-dimethylphenol	1,27	3,41	0,109	201,07	32,55	0,395	0,166	9,10
110	4-Bromo-6-chloro-2-cresol	1,28	3,46	-0,226	221,48	33,75	0,393	0,167	9,18

111	4-Cyclopentylphenol	1,29	3,44	0,437	162,25	30,74	0,405	0,160	9,10
112	2-*Tert*-butylphenol	1,29	3,17	0,436	150,24	31,61	0,394	0,166	9,45
113	2-*Tert*-butyl-4-methylphenol	1,30	3,63	0,477	164,27	31,02	0,394	0,166	9,57
114	2-Hydroxydiphenylmethane	1,31	3,47	0,242	184,25	38,66	0,360	0,217	9,31
115	Butyl-4-hydroxybenzoate	1,33	3,41	-0,367	194,25	35,89	1,083	0,221	9,37
116	3-Phenylphenol	1,35	3,23	0,161	170,22	40,68	0,360	0,217	9,27
117	*n*-Pentyloxyphenol	1,36	3,43	0,330	180,27	29,37	0,732	0,219	9,25
118	2,4-Dibromophenol	1,40	3,31	-0,349	251,9	31,45	0,397	0,164	8,98
119	2,4,6-Trichlorophenol	1,41	2,75	-0,502	197,44	21,69	0,385	0,171	9,01
120	2-Hydroxy-4-methoxybenzophenone	1,42	3,43	-0,574	228,26	43,71	1,196	0,172	9,75
121	Isoamyl-4-hydroxybenzoate	1,48	3,76	-0,363	208,28	34,01	1,083	0,221	9,57
122	3,5-Dichlorosalicylaldehyde	1,55	2,41	-0,893	191,01	27,49	0,742	0,173	9,10
123	4-Cyclohexylphenol	1,56	4,00	0,442	176,28	29,22	0,360	0,217	9,14
124	3,5-Dichlorophenol	1,57	3,25	-0,285	163,00	25,71	0,390	0,169	8,82
125	3,5-Di-*tert*-butylphenol	1,64	4,86	0,470	206,36	24,80	0,390	0,169	9,72
126	3,5-Dibromosalicylaldehyde	1,64	2,67	-0,924	279,91	31,64	0,821	0,174	9,22
127	3,4-Dichlorophenol	1,75	3,19	-0,236	163,00	29,66	0,390	0,169	8,79
128	4-Bromo-2,6-dichlorophenol	1,78	2,69	-0,514	241,89	25,89	0,389	0,169	9,05
129	2,6-Di-*tert*-butyl-4-methylphenol	1,80	5,32	0,383	220,39	27,73	0,359	0,217	10,38
130	4-Chloro-2-isopropyl-5-methyl-phenol	1,85	4,22	0,114	184,68	28,98	0,394	0,166	9,53
131	2,4,6-Tribromophenol	2,03	3,28	-0,621	330,79	27,92	0,399	0,162	9,22
132	4-Heptyloxyphenol	2,03	4,50	0,329	208,33	27,82	0,732	0,219	9,07
133	4-*Tert*-octylphenol	2,10	4,93	0,474	206,36	25,02	0,360	0,217	9,26
134	4-(4-Bromophenyl)phenol	2,31	3,95	-0,399	249,11	36,69	0,360	0,217	9,12
135	3,5-Diiododsalicylaldehyde	2,34	2,90	-0,901	373,91	34,31	0,708	0,194	9,28
136	2,3,5-Trichlorophenol	2,37	2,84	-0,578	197,44	22,87	0,350	0,220	8,98
137	4-Nonylphenol	2,47	6,19	0,429	220,39	24,65	0,360	0,217	9,15
138	Nonyl-4-hydroxybenzoate	2,63	6,07	-0,368	264,40	29,71	1,083	0,221	9,12

139	2,4,6-Trinitrophenol	-0,16	-4,98	-2,534	229,12	30,49	0,359	0,217	9,13
140	3,4-Dinitrophenol	0,27	0,24	-1,863	184,12	33,92	0,386	0,172	8,84
141	2,6-Dinitrophenol	0,54	-1,67	-1,952	184,12	31,00	0,393	0,166	9,05
142	2,6-Dichloro-4-nitrophenol	0,63	-0,66	-1,441	208,00	16,45	0,388	0,169	9,03
143	2,5-Dinitrophenol	0,95	-0,16	-2,262	184,12	25,53	0,385	0,172	8,96
144	2,4-Dinitrophenol	1,08	-1,57	-1,887	184,12	27,64	0,359	0,217	8,92
145	2,6-Dinitro-4-cresol	1,23	-0,87	-1,893	198,15	31,03	0,359	0,217	9,17
146	4-Bromo-2-fluoro-6-nitrophenol	1,62	0,13	-1,650	236,00	19,04	0,354	0,218	9,52
147	Pentafluorophenol	1,64	0,80	-1,296	184,07	1,90	0,370	0,181	8,30
148	4,6-Dinitro-2-methylphenol	1,72	-0,73	-1,825	198,15	27,5	0,359	0,217	9,21
149	2,4-Dichloro-6-nitrophenol	1,75	0,70	-1,579	208,00	15,72	0,384	0,172	9,06
150	Pentachlorophenol	2,05	2,11	-0,978	266,32	17,97	0,381	0,175	9,20
151	2,3,5,6-Tetrachlorophenol	2,22	1,80	-0,817	231,88	21,35	0,383	0,173	9,14
152	Pentabromophenol	2,66	3,18	-1,193	488,57	27,07	0,403	0,16	9,52
153	2,3,4,5-Tetrachlorophenol	2,71	3,16	-0,752	231,88	21,36	0,384	0,173	9,05
154	4-Acetamidophenol	-0,82	0,34	0,253	151,18	39,49	1,024	0,218	8,88
155	3-Aminophenol	-0,52	0,34	0,522	109,14	46,32	0,684	0,162	8,73
156	4-Aminophenol	-0,08	-0,29	0,439	109,14	42,89	0,684	0,162	8,70
157	3-Methylcatechol	0,28	1,34	0,268	124,15	39,79	0,744	0,219	17,81
158	2-Amino-4-*tert*-butylphenol	0,37	2,13	0,418	165,26	29,75	0,677	0,164	9,18
159	4-Methylcatechol	0,37	1,34	0,332	124,15	41,10	0,744	0,219	17,64
160	1,2,4-Trihydroxybenzene	0,44	0,06	0,133	126,12	48,00	1,116	0,220	26,04
161	Hydroquinone	0,47	0,64	0,233	110,12	45,14	0,748	0,219	17,29
162	Catechol	0,75	0,88	0,297	110,12	42,31	0,744	0,219	17,34
163	2-Amino-4-chlorophenol	0,78	1,67	0,043	143,58	34,89	0,681	0,162	8,86
164	1,2,3-Trihydroxybenzene	0,85	0,28	0,029	126,12	45,53	1,113	0,220	26,09
165	2-Aminophenol	0,94	2,44	0,406	109,14	38,19	0,681	0,162	8,79
166	4-Chlorocatechol	1,06	2,13	0,001	144,56	36,19	0,738	0,220	17,5

167	Chlorohydroquinone	1,26	1,51	-0,111	144,56	40,15	0,740	0,220	17,54
168	4-Amino-2-cresol	1,31	0,17	0,413	123,17	39,15	0,654	0,216	8,99
169	2,3-Dimethylhydroquinone	1,41	1,56	0,215	138,18	40,41	0,747	0,219	18,23
170	4-Amino-2,3-dimethylphenol	1,44	0,63	0,406	137,20	39,25	0,681	0,163	9,17
171	Bromohydroquinone	1,68	2,00	-0,186	189,01	38,80	0,752	0,219	17,68
172	Tetrachlorocatechol	1,70	3,07	-0,830	247,88	21,27	0,723	0,221	18,17
173	Phenylhydroquinone	2,00	2,09	-0,229	186,22	44,77	0,750	0,219	18,86
174	3,5-Di-*tert*-butylcatechol	2,11	4,26	0,294	222,36	28,05	0,744	0,219	19,63
175	Methoxyhydroquinone	2,20	0,47	0,226	140,15	46,98	1,103	0,220	17,88
176	3-Hydroxy-4-nitrobenzaldehyde	0,27	0,43	-1,755	167,13	26,22	0,751	0,168	8,99
177	5-Hydroxy-2-nitrobenzaldehyde	0,33	0,65	-1,486	167,13	29,79	0,787	0,186	8,87
178	2-Amino-4-nitrophenol	0,47	0,59	-1,116	154,14	30,05	0,651	0,217	8,88
179	4-Methyl-2-nitrophenol	0,57	1,92	-1,141	153,15	25,90	0,359	0,217	8,96
180	4-Hydroxy-3-nitrobenzaldehyde	0,61	-0,36	-1,456	167,13	26,28	0,751	0,168	8,94
181	4-Nitrosophenol	0,65	0,51	-0,796	123,12	41,47	0,839	0,182	8,71
182	2-Nitroresorcinol	0,66	-0,98	-1,321	155,12	21,13	0,747	0,219	17,72
183	4-Methyl-3-nitrophenol	0,74	2,37	-1,109	153,15	28,46	0,393	0,167	8,88
184	2-Chloromethyl-4-nitrophenol	0,75	0,73	-1,195	187,59	28,47	0,393	0,167	9,11
185	2-Bromo-2^0-hydroxy-5^0-nitroacetanilide	0,87	0,71	-1,105	275,07	22,41	1,041	0,232	9,21
186	4-Amino-2-nitrophenol	0,88	0,53	-1,120	154,14	29,59	0,683	0,162	8,91
187	2-Fluoro-4-nitrophenol	1,07	0,01	-1,333	157,11	24,13	0,353	0,219	9,01
188	5-Fluoro-2-nitrophenol	1,13	0,76	-1,447	157,11	19,23	0,386	0,172	8,78
189	4-Nitrocatechol	1,17	1,05	-1,160	155,12	31,31	0,744	0,219	17,55
190	2-Amino-4-chloro-5-nitrophenol	1,17	2,38	-0,960	188,58	28,76	0,681	0,162	8,99
191	4-Fluoro-2-nitrophenol	1,38	1,21	-1,447	157,11	25,27	0,354	0,219	8,93
192	4-Nitrophenol	1,42	1,21	-1,065	139,12	26,69	0,394	0,166	8,72
193	2-Chloro-4-nitrophenol	1,59	0,30	-1,264	173,56	23,75	0,393	0,166	8,87
194	4-Chloro-6-nitro-3-cresol	1,64	2,31	-1,346	187,59	25,47	0,394	0,166	9,09

195	3-Methyl-4-nitrophenol	1,73	1,74	-1,007	153,15	26,54	0,360	0,217	8,90
196	4-Bromo-2-nitrophenol	1,87	1,41	-1,398	218,01	30,38	0,361	0,217	9,04
197	4-Chloro-2-nitrophenol	2,05	1,68	-1,388	173,56	18,73	0,394	0,166	8,91

Table 2.2. Chemicals, their toxicity to *Tetrahymena pyriforms* and significant physico-chemical descriptors (testing subset)

N	Compound	Toxicity	Log *D*	LUMO	MW	P_{NEG}	ABSQon	MaxHp	SsOH
1	2-Hydroxybenzylalcohol	-0,95	0,30	0,344	124,15	37,22	0,783	0,210	17,509
2	2-Fluorophenol	0,19	1,69	0,013	112,11	45,03	0,392	0,166	8,544
3	2-Allylphenol	0,33	2,50	0,348	134,19	39,70	0,394	0,166	9,194
4	3-Chlorophenol	0,87	2,39	0,019	128,56	31,55	0,394	0,166	8,729
5	4,6-Dichlororesorcinol	0,97	2,37	-0,263	179,00	29,48	0,734	0,220	17,748
6	4-Benzyloxyphenol	1,04	2,96	0,232	200,25	40,93	0,728	0,166	9,086
7	3-Iodophenol	1,12	2,92	-0,070	220,01	35,87	0,394	0,166	8,808
8	2,3-Dichlorophenol	1,28	2,61	-0,262	163,00	31,13	0,389	0,169	8,884
9	4-Phenylphenol	1,39	3,20	-0,086	170,22	41,48	0,360	0,217	9,104
10	4-Hexyloxyphenol	1,64	3,97	0,330	194,30	28,51	0,732	0,219	9,043
11	4-Hexylresorcinol	1,80	3,88	0,327	194,30	27,84	0,748	0,219	18,575
12	2,4,5-Trichlorophenol	2,10	3,27	-0,555	197,44	25,58	0,386	0,171	8,950
13	2-Ethylhexyl-4^0-hydroxybenzoate	2,51	5,34	-0,366	250,37	32,10	1,011	0,160	9,121
14	2,3,5,6-Tetrafluorophenol	1,17	0,63	-0,994	166,08	13,28	0,367	0,183	8,338
15	3,4,5,6-Tetrabromo-2-cresol	2,57	4,69	-0,882	423,71	30,72	0,402	0,161	9,565
16	Trimethylhydroquinone	1,34	2,02	0,215	152,21	38,88	0,747	0,219	18,695
17	4-Nitro-3-(trifluoromethyl)-phenol	1,65	2,13	-1,585	207,12	33,09	0,360	0,217	8,760
18	4-Ethoxyphenol	0,01	1,84	0,327	138,18	40,58	0,756	0,172	8,869
19	4-Bromo-2,6-dimethylphenol	1,17	3,41	0,085	201,07	34,29	0,361	0,217	9,310
20	4-Methoxyphenol	-0,14	1,31	0,303	124,15	46,44	0,759	0,172	8,797

2.2. Proposed in literature approaches which have been used for comparison with the artificial neural networks

Mark Cronin et al. in [5] have proposed and compared different approaches for developing of QSARs for the prediction of the toxicity of 200 phenols to Tetrahymena pyriformis based on 108 physico-chemical descriptors. Among them are response-surface approach or two parameters approach, stepwise regression or seven parameters approach, two dimension and three dimension partial least squares. Disadvantages of presented in [5] approaches are observation of outliers. But in general all these approaches were found to be a good model for solving the task of prediction of the phenols toxicity.

1. Response-surface analyses QSAR (or two-parameter QSAR)

An effort to model the complete data set using the response-surface approach was applied. The following relationship was found between the toxicity of the phenols to T. pyriformis and log *D* and E_{LUMO} (log D was found to be more successful in modelling toxicity than log *P*):

$$log(IGC_{50})^{-1} = 0{,}53(0{,}022)logD - 0{,}96(0{,}048)E_{LUMO} - 0{,}58(0{,}057) \quad (2.2)$$

2. Stepwise regression analysis QSAR (or seven-parameter QSAR)

Seven descriptors were identified as being important to describe toxicity, and no redundancy was observed between them.

$$log(IGC_{50})^{-1} = 0{,}38(0{,}024)logD - 0{,}58(0{,}058)E_{LUMO} + 0{,}0047(0{,}0008)MW - 0{,}018(0{,}0048)P_{NEG} + 0{,}050(0{,}0083)SsOH - 0{,}61(0{,}11)ABSQon + 2{,}69(1{,}15)MaxHp - 0{,}99(0{,}29) \quad (2.3)$$

3. Regression on partial least squares

Thus, in order to reduce the number of descriptors, principal component analysis (PCA) was performed. PCA indicated that are seven major components

describing over 80% of the total variance on the 108 descriptors, from these components 14 individual variables were chosen to represent the full descriptor set.

In order to investigate the possibility of improving the modelling by the selection of relevant descriptors, PLS was repeated utilising those variables found to be useful in both the response-surface and stepwise regression analyses. In total there are 11 descriptors.

So, in [5] were used two dimension and three dimension PLS models.

2.3. Describing of the used artificial neural networks and choosing of their parameters

The software package MATLAB R2021b Update 2 (9.11.0.1837725) along with the Neural Network Toolbox and Statistical Toolbox were used in the present work (license 10232054 trial – individual) [44−46].

2.3.1. Feed-forward and cascade neural networks

Feed-forward networks as any artificial neural network consist of layers. The signals are sent to the input layer. Than they are processed by neurons of the hidden layer. And outputs (final result) are produced by output layer. So, the signals moves in only direct direction. Each subsequent layer has a connection from the previous layer.

Figure 2.1 shows a single-layer feed-forward network [46, p. 136].

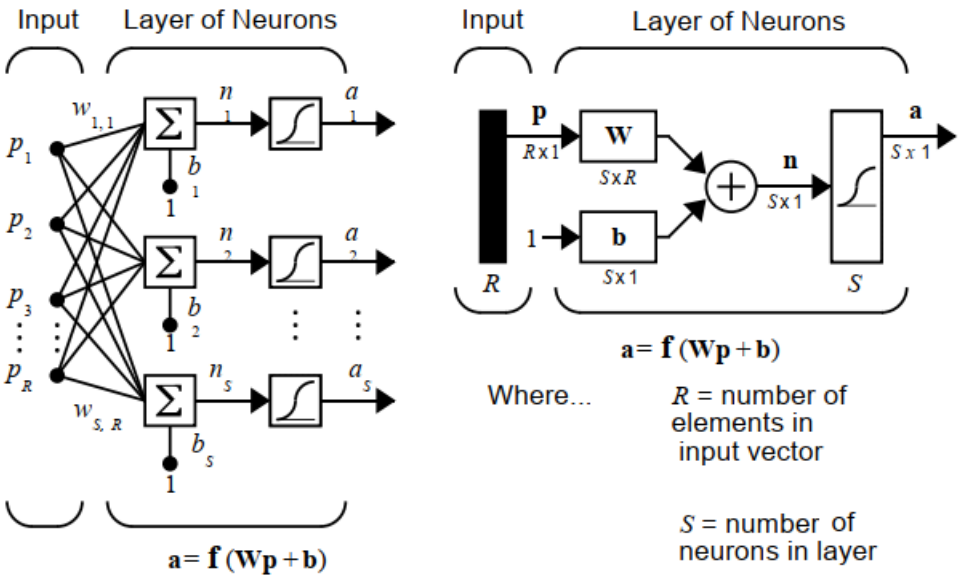

Figure 2.1. Feed-forward neural network architecture [46, p. 136]

Cascade forward network is a variation on the feed-forward network but with one difference. Cascade forward network has additional connections from the input to every layer, and from each layer to all following layers.

Feed-forward networks and cascade forward network with one hidden layer and enough neurons in the hidden layer can be used for processing of any information and for solving various tasks.

Feed-forward and cascade networks as usually have one (rarely more than one) hidden layer of neurons with sigmoid transfer function followed by an output layer of neurons with linear transfer function. Presenting of nonlinear transfer functions in artificial neural network architecture allow to learn nonlinear and linear input-output relationships.

For correct application of artificial neural network the transfer or activation functions in each layer, the learning rule, and the number of neurons in each layer for constructing the architecture of an artificial neural network should be determined. The input signals multiplied by the weight parameters are summed and passed through a transfer function to produce the output for neurons.

The initial weights were assigned throughout according to the Nguyen-Widrow method. This method is fast and in combination with the Levenberg-

Marquardt training method provides rapid convergence and accuracy of the ANN algorithms as a whole. Besides, the Nguyen-Widrow algorithm is best suited for use with the sigmoid and linear transfer functions. Different numbers of hidden neurons were examined to choose the most efficient architecture of feed-forward and cascade neural networks [21, 47−49]. In [21] it was shown that for feed-forward and cascade neural networks the best results were attained with the Levenberg-Marquardt training method and tangent sigmoid / linear for hidden and output layers.

Table 2.3. Training parameters of feed-forward and cascade neural networks

Type of parameter	Description
Learning type	Supervision
Training methods	Levenberg-Marquardt
Minimized error function	Mean squared error
Number of hidden neurons	*n*−*m*
Transfer functions	Tangent sigmoid and linear
Maximal possible number of training epochs	500 (has never been reached.)
Learning method	Gradient optimization algorithm
Initialization method	Nguyen-Widrow algorithm

Two of the most commonly used transfer functions were tested, namely linear [Eq. (2.4)], and tangent sigmoid [Eq. (2.5)]:

$$f_{\text{linear}} = n, \tag{2.4}$$

$$f_{\text{tangent sigmoid}} = \frac{e^{n} - e^{-n}}{e^{n} + e^{-n}}, \tag{2.5}$$

$$n = wp + b, \tag{2.6}$$

where p is the input vector, w is the weight vector of neuron, and b is the bias of neuron.

The mean squared error was used as the error function to be minimized during the ANN training [48]:

$$MSE = \frac{1}{M}\sum_{i=1}^{M}(y_i - t_i)^2 , \qquad (2.7)$$

where M is the number of phenols in the training subset, y_i and t_i are the predicted and real outputs of the i-th phenol, correspondingly.

To improve the training efficiency of ANN, the gradient descent optimization was used as an additional method.

The number of hidden neurons is optimal if the neural network is trained properly and correct predict the phenols toxicities from the testing subset. If the number of hidden neurons is small, the performance of algorithms may be unsatisfactory. If the number of hidden neurons is too high, over-fitting is expected.

We have found the dependencies of the network training performance on the numbers of hidden neurons for the feed-forward and cascade neural networks (Figures 2.2 and 2.3).

So, the optimal number of hidden neurons for feed-forward neural network is nine, the optimal number of hidden neurons for cascade neural network is six. The minimum values of mean squared errors of neural networks training are observed at this numbers of hidden neurons.

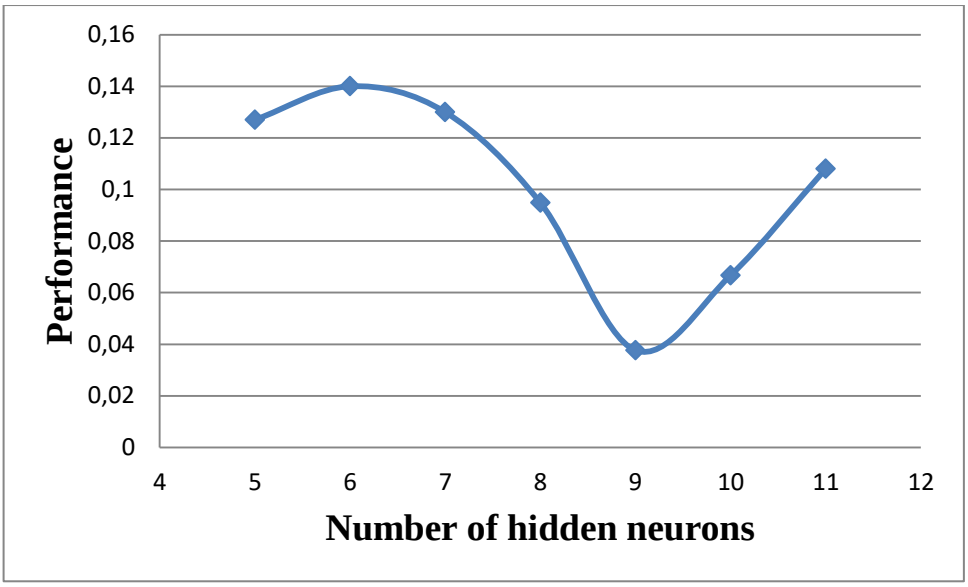

Figure 2.2. Dependence of mean squared error on the numbers of hidden neurons for feed-forward neural network

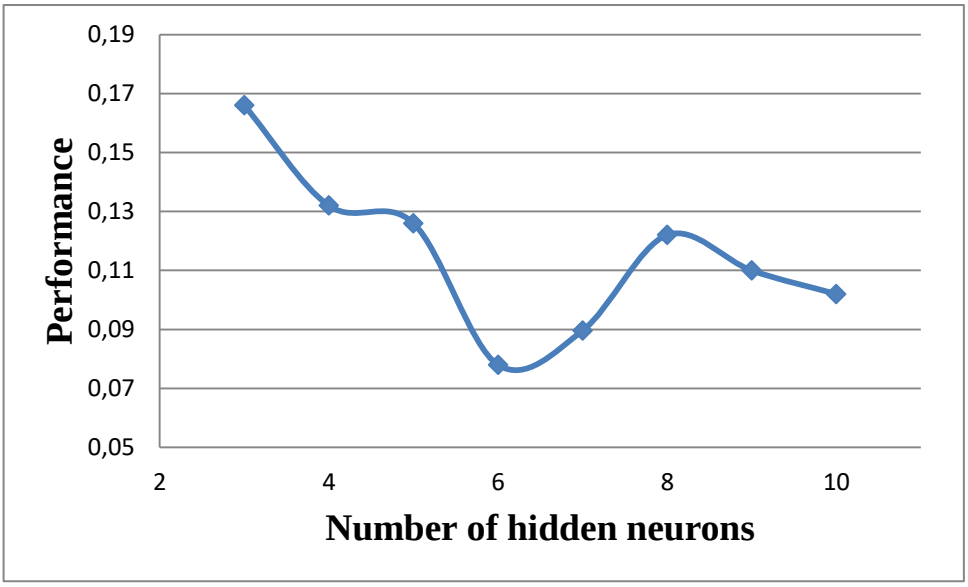

Figure 2.3. Dependence of mean squared error on the numbers of hidden neurons for cascade neural network

2.3.2. The radial basis function networks

Radial basis function network also consists of an input layer, a hidden layer and an output layer. Each of these layers has different roles. The input layer include the same number of neurons as predictor variables. The role of input neurons is transferring input information to the hidden layer. The input neurons does not process the input vectors.

The transformation of the vector of characteristics from the input layer to the hidden layer is nonlinear, while the transformation from the hidden layer to the output layer is linear. Every hidden neuron can store only one element of the training set, so number of hidden neurons as many as number of samples in training set (input vectors).

Layer of hidden neurons uses a radial basis function as activation nonlinear function for processing the input vectors.

The output of the each hidden neuron of this type of artificial neural networks can be written as:

$$q_i = g_i(\|X - C_i\|) = exp\left(-\frac{\|X-C_i\|^2}{2\sigma_i^2}\right), \quad (2.8)$$

where $X = (x_1, x_2, \ldots, x_m)$ is an m-dimensional vector; $g_i(X)$ is a Gaussian activation function, $i=1,2,\ldots,n$; n is the number of neurons in the hidden layer; C_i is the center of the i-th activation function; $\|*\|$ is the Euclid norm; and σ_i is the width of the receptive field (deviation of the activation function or spread parameter).

The activation of the output layer is a linear combination of the units in the hidden layer elements, which can be expressed as:

$$y = \sum_{i=1}^{n} w_i q_i, \quad (2.9)$$

where w_i are weight coefficients connecting the hidden layer to the output layer [24, 26, 50]. For prediction, there is one output neuron for each sample in training / testing set. Hidden and output layers have biases.

So, one can see that the design of radial basis function networks involves selecting centers, spread parameter, biases and weights. Each bias in the hidden layer is set to 0.8326/spread. The centers of activation function are selected randomly from the data [21].

Two types of radial basis function neural networks were used in this work:

1) RBFN − a radial basis function network, in architecture of which one neuron are added at a time. The process of neurons addition to the network is continued until the error goal reached its given value or a maximum number of neurons has been reached;

2) RBEFN − a radial basis function network with zero error on training vectors.

In the case of both types of radial basis function neural networks it is important to choose correct value of spread parameter. It should be large enough that the radial basis neurons were able to distinguish overlapping regions of the input vectors, and not so large that all the radial basis neurons form the same outputs [46].

Figure 2.4 shows general scheme of radial basis network and Figure 2.5 shows detailed radial basis network architecture [46, p. 247, 248].

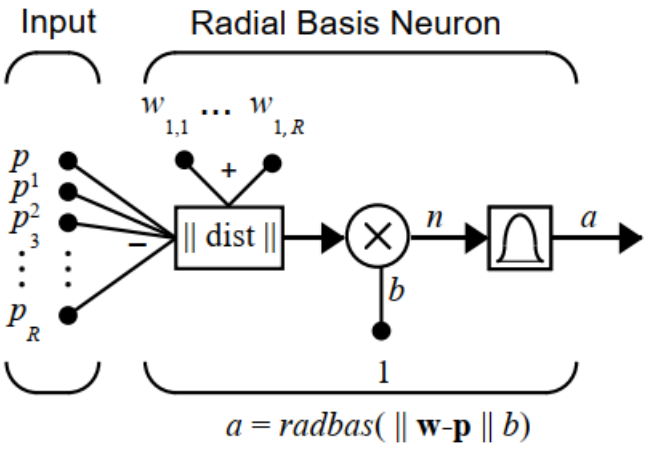

Figure 2.4. General scheme of radial basis network [46, p. 247]

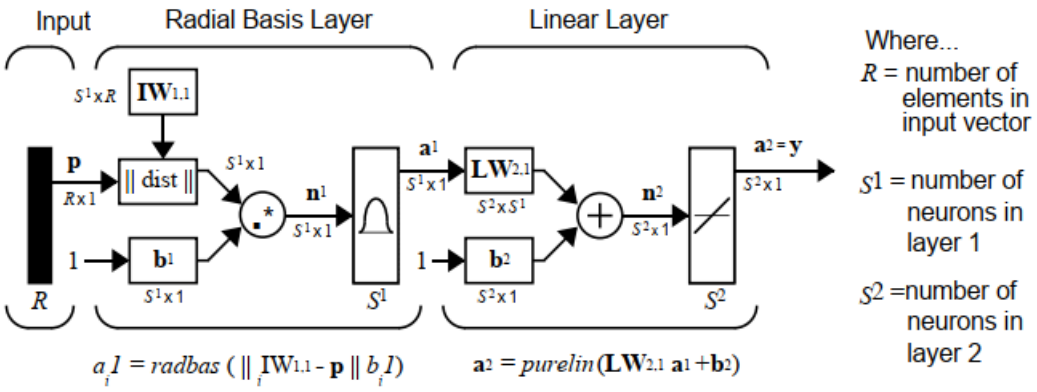

Figure 2.5. Radial basis network architecture [46, p. 248]

As stated above it is important to choose optimal value of the spread parameter. That is why we have found the optimal value of spread of radial basis function. Two types of radial basis networks were trained with zero error. That is why for choosing correct value of parameter spread we have calculated mean squared error for testing subset [Eq. (2.7)].

The values of mean squared error for testing set for RBFN and RBEFN at different spread values are shown in Table 2.4.

Table 2.4. The values of mean squared error for testing subset of compounds for radial basis function networks at different spread values

Spread value	Mean squared error	
	RBFN	RBEFN
1,0	1,70	3,37
0,9	2,80	2,80
0,8	2,48	2,48
0,7	2,32	2,32
0,6	306,80	306,80

So, the minimum mean squared error is observed at spread 1,0 for RBFN and at spread 0,7 for RBFEN.

Conclusions to section 2

1. The applicability of the feed-forward, cascade and radial basis artificial neural networks for the prediction of toxicity of phenols on the basis of their molecular descriptors has been explored.

2. The optimal architectures (the number of hidden neurons, spread value) for the feed-forward, cascade and radial basis artificial neural networks have been determined.

SECTION 3. RESULTS AND DISCUSSION

This study describes the comparative development and validation of ANNs for the toxicity to Tetrahymena pyriformis of a heterogeneous group of phenols utilising a large data set of ubiquitous and easily calculated descriptors.

Results obtained by using of four artificial neural networks were compared with the results of proposed in literature approaches [5]. These approaches are described in section 2.2.

Plot of observed toxicity for the testing subset of compounds against that predicted from different artificial neural networks (with chosen optimal number of hidden neurons and spread value) are presented on Figures 3.1−3.4.

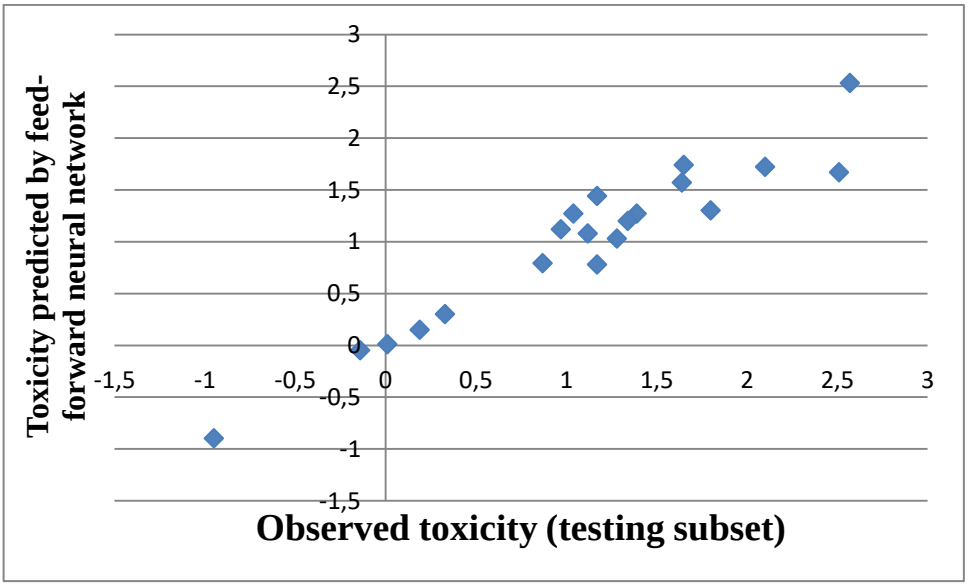

Figure 3.1. Plot of observed toxicity for the testing subset of compounds against that predicted from feed-forward neural network (number of hidden neurons = 9)

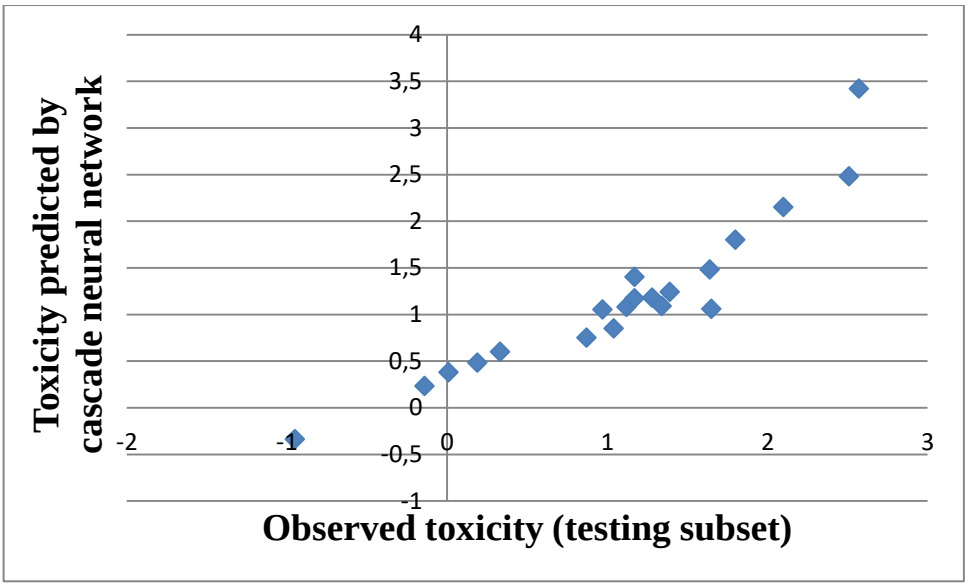

Figure 3.2. Plot of observed toxicity for the testing subset of compounds against that predicted from cascade neural network (number of hidden neurons = 6)

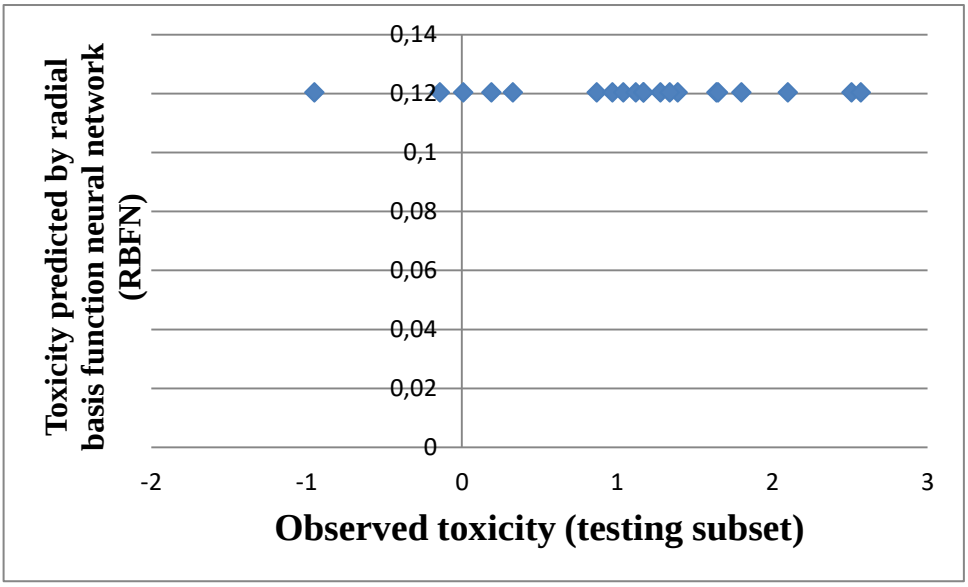

Figure 3.3. Plot of observed toxicity for the testing subset of compounds against that predicted from radial basis function neural network (RBFN) (spread value = 1,0)

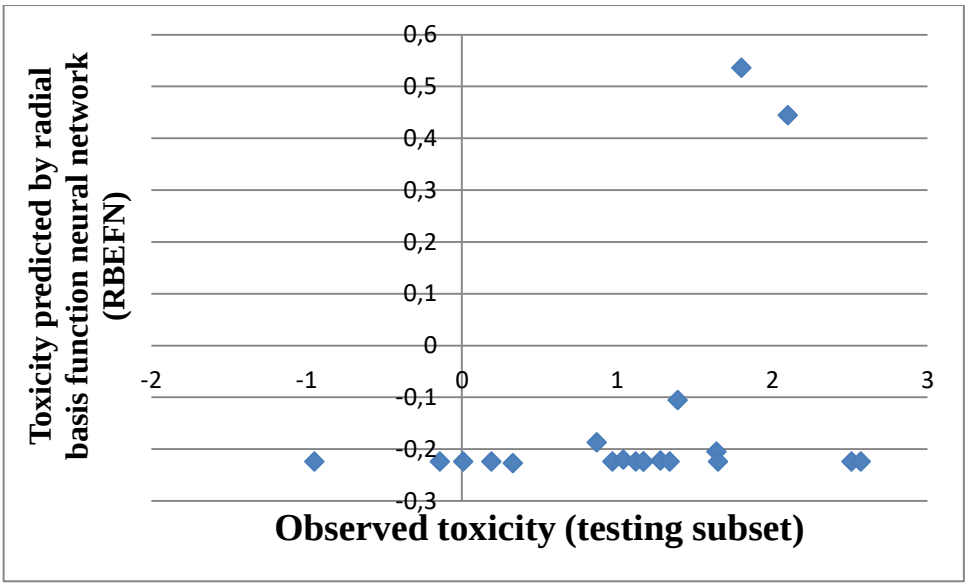

Figure 3.4. Plot of observed toxicity for the testing subset of compounds against that predicted from radial basis function neural network (RBEFN) (spread value = 0,7)

So, results of prediction of phenols toxicity by using of both radial basis function neural networks are bad. RBFN predicted for all 20 compounds of testing set the same value of toxicity (-0,1204). RBEFN predicted for all 20 compounds of testing set the different value of toxicity, but these results are characterized by large deviations form observed toxicity. And also RBEFN for compounds with positive toxicity predicted negative toxicity and vise verse.

Results obtained by using of radial basis function networks are worse that results obtained by using of feed-forward and cascade neural networks. A few points need to be noted.

Training a radial basis function network often takes much less time than training any network with sigmoid transfer function for hidden layer and linear transfer function for output layer. Radial basis function networks usually have many times more neurons than a cascade neural network or feed-forward network with the using of sigmoid transfer function in the hidden layer. This situation can be explained by the fact that neurons with sigmoid transfer function can have outputs over a large space of the input information, while neurons with radial basis

transfer function only respond to relatively small regions of the input information [46].

The values of mean squared error [Eq. (2.7)] for testing subset of compounds for used in this work artificial neural networks (with chosen optimal number of hidden neurons and spread value) and proposed in literature [5] methods are shown in Table 3.1.

Table 3.1. The values of mean squared error for testing subset of compounds used in this work artificial neural networks (with chosen optimal number of hidden neurons and spread value) and proposed in literature methods

Method	Mean squared error
Feed-forward neural network	0,0766
Cascade neural network	0,105
Radial basis function neural network (RBFN)	1,704
Radial basis function neural network (RBEFN)	2,319
Two-parameter QSAR [Eq. (2.2)]	0,131
Seven-parameter QSAR [Eq. (2.3)]	0,145
Two dimension PLS model	0,180
Three dimension PLS model	0,148

We can make conclusion, that feed-forward and cascade neural networks are more effective for prediction of phenols toxicity than proposed in literature [5] methods.

Predicted toxicities and residual values for the testing subset of compounds obtained by using of feed-forward neural network, cascade neural network and proposed in literature methods are shown in Table 3.2.

Table. 3.2. Predicted toxicities and residuals for the testing subset of compounds

N	Compound	Obser-ved toxicity	Two-parameter QSAR [5]		Seven-parameter QSAR [5]		Two dimension PLS model [5]		Three dimension PLS model [5]		Feed-forward neural network (number of hidden neurons = 9)		Cascade neural network (number of hidden neurons = 6)	
			Pred[a]	Res[b]	Pred[a]	Res[b]	Pred[a]	Res[b]	Pred[a]	Res[b]	Pred[a]	Res[b]	Pred[a]	Res[b]
1	2-Hydroxybenzylalcohol	-0,95	-0,75	-0,20	-0,19	-0,77	-0,08	-1,03	-0,41	-0,54	-0,90	-0,05	-0,34	-0,61
2	2-Fluorophenol	0,19	0,30	-0,11	0,00	0,19	0,04	0,14	-0,10	0,29	0,15	0,04	0,48	-0,29
3	2-Allylphenol	0,33	0,40	-0,07	0,34	-0,01	0,28	0,05	0,32	0,01	0,30	0,03	0,60	-0,27
4	3-Chlorophenol	0,87	0,66	0,21	0,59	0,29	0,60	0,27	0,54	0,33	0,79	0,08	0,75	0,12
5	4,6-Dichlororesorcinol	0,97	0,92	0,05	1,41	-0,44	1,14	-0,17	0,86	0,10	1,12	-0,15	1,05	-0,08
6	4-Benzyloxyphenol	1,04	0,76	0,28	0,66	0,38	0,63	0,40	0,58	0,46	1,27	-0,23	0,85	0,19
7	3-Iodophenol	1,12	1,03	0,09	1,19	-0,07	0,83	0,29	1,06	0,06	1,08	0,04	1,08	0,04
8	2,3-Dichlorophenol	1,28	1,05	0,23	1,02	0,26	0,99	0,28	0,86	0,42	1,03	0,25	1,18	0,10
9	4-Phenylphenol	1,39	1,19	0,20	1,15	0,25	0,82	0,57	0,89	0,50	1,27	0,12	1,24	0,15
10	4-Hexyloxyphenol	1,64	1,20	0,44	1,31	0,33	1,13	0,33	1,28	0,36	1,57	0,07	1,48	0,16
11	4-Hexylresorcinol	1,80	1,15	0,65	1,76	0,04	1,38	0,41	1,30	0,50	1,30	0,50	1,80	0,00
12	2,4,5-Trichlorophenol	2,10	1,68	0,42	1,71	0,39	1,66	0,44	1,45	0,65	1,72	0,38	2,15	0,05
13	2-Ethylhexyl-4^0-hydroxybenzoate	2,51	2,59	-0,08	2,11	0,40	1,89	0,62	1,72	0,79	1,67	0,84	2,48	0,03
14	2,3,5,6-Tetrafluorophenol	1,17	0,71	0,46	1,06	0,11	1,49	-0,33	0,95	0,21	0,78	0,39	1,17	0,00

15	3,4,5,6-Tetrabromo-2-cresol	2,57	2,74	-0,17	3,40	-0,82	2,86	-0,28	2,91	-0,33	2,53	0,04	3,42	-0,85
16	Trimethylhydroquinone	1,34	0,28	1,06	0,74	0,60	0,43	0,91	1,27	0,08	1,20	0,14	1,09	0,15
17	4-Nitro-3-(trifluoromethyl)-phenol	1,65	2,07	-0,42	1,92	-0,27	1,57	0,08	1,30	0,35	1,74	-0,09	1,06	0,59
18	4-Ethoxyphenol	0,01	0,08	-0,07	-0,11	0,12	0,07	-0,06	-0,08	0,09	0,01	0,00	0,38	-0,37
19	4-Bromo-2,6-dimethylphenol	1,17	1,14	0,03	1,41	-0,24	1,12	0,05	1,26	-0,09	1,44	-0,27	1,40	-0,27
20	4-Methoxyphenol	-0,14	-0,18	0,04	-0,47	0,33	-0,23	0,09	-0,45	0,31	-0,05	-0,09	0,23	-0,37

[a]Predicted toxicity

[b]Residual (observed toxicity – predicted toxicity)

Conclusions to section 3

1. Radial basis function neural networks are inapplicable for prediction of phenols toxicity.

2. The efficiency of prediction algorithms decreases in that order: feed-forward neural network > cascade neural network > response-surface analyses [Eq. (2.2)] > stepwise regression analysis [Eq. (2.3)] > three dimension PLS model > two dimension PLS model.

CONCLUSIONS

1. The applicability of the feed-forward, cascade and radial basis function neural networks for the prediction of toxicity of phenols on the basis of their seven molecular descriptors has been explored, and the optimal parameters of the ANNs which provide the correct prediction have been determined.

2. The feed-forward and cascade neural networks are the most suitable for the prediction of toxicity of phenols and sufficiently exceed the commonly used methods (QSARs and PLSs). In contrast, radial basis function neural network gave poor agreement with experiment (observed toxicity).

3. The ability to obtain the toxicity of phenols theoretically without performing additional experiments provides a valuable tool that can be utilized in practice in creation of medicines. And it will contribute to the further development of computational methods for predicting the toxicity of chemical compounds.

REFERENCES

1. Use of QSAR methodology in toxicology / L. S. Bobkova et al. *Современные проблемы токсикологии.* 2008. Vol. 2. P. 78−86. [in Ukrainian].

2. Flavonoids and Other Phenolic Compounds from Medicinal Plants for Pharmaceutical and Medical Aspects: An Overview / D. Tungmunnithum et al. *Medicines*. 2018. Vol. 5, no. 3. P. 93−109. URL: https://doi.org/10.3390/medicines5030093.

3. Role of Phenolic Compounds in Human Disease: Current Knowledge and Future Prospects / M. M. Rahman et al. *Molecules*. 2021. Vol. 27, no. 1. P. 233−269. URL: https://doi.org/10.3390/molecules27010233

4. Mutha R. E., Tatiya A. U., Surana S. J. Flavonoids as natural phenolic compounds and their role in therapeutics: an overview. *Future journal of pharmaceutical sciences*. 2021. Vol. 7, no. 1. P. 1−13. URL: https://doi.org/10.1186/s43094-020-00161-8

5. Comparative assessment of methods to develop QSARs for the prediction of the toxicity of phenols to Tetrahymena pyriformis / M. T. Cronin et al. *Chemosphere*. 2002. Vol. 49, no. 10. P. 1201−1221.

6. Cronin M. T. D., Schultz T. W. Structure-toxicity relationships for phenols to Tetrahymena pyriformis. *Chemosphere*. 1996. Vol. 32, no. 8. P. 1453–1468. URL: https://doi.org/10.1016/0045-6535(96)00054-9

7. Garg R., Kurup A., Hansch C. Comparative QSAR: On the Toxicology of the Phenolic OH Moiety. *Critical Reviews in Toxicology*. 2001. Vol. 31, no. 2. P. 223–245. URL: https://doi.org/10.1080/20014091111686

8. Multivariate Discrimination between Modes of Toxic Action of Phenols / A. Aptula et al. *Quantitative Structure-Activity Relationships*. 2002. Vol. 21, no. 1. P. 12−22. URL: https://doi.org/10.1002/1521-3838(200205)21:1%3C12::aid-qsar12%3E3.0.co;2-m

9. Prediction of acute toxicity of phenol derivatives using multiple linear regression approach for Tetrahymena pyriformis contaminant identification in a

median-size database / K. Dieguez-Santana et al. *Chemosphere*. 2016. Vol. 165. P. 434–441. URL: https://doi.org/10.1016/j.chemosphere.2016.09.041

10. Abbasitabar F., Zare-Shahabadi V. In silico prediction of toxicity of phenols to Tetrahymena pyriformis by using genetic algorithm and decision tree-based modeling approach. *Chemosphere*. 2017. Vol. 172. P. 249–259. URL: https://doi.org/10.1016/j.chemosphere.2016.12.095

11. In silico prediction of toxic action mechanisms of phenols for imbalanced data with Random Forest learner / J. Chen et al. *Journal of Molecular Graphics and Modelling*. 2012. Vol. 35. P. 21–27. URL: https://doi.org/10.1016/j.jmgm.2012.01.002

12. Ren S. Phenol mechanism of toxic action classification and prediction: a decision tree approach. *Toxicology Letters*. 2003. Vol. 144, no. 3. P. 313–323. URL: https://doi.org/10.1016/s0378-4274(03)00236-4

13. Polyphenols: A concise overview on the chemistry, occurrence, and human health / A. Durazzo et al. *Phytotherapy Research*. 2019. Vol. 33, no. 9. P. 2221–2243. URL: https://doi.org/10.1002/ptr.6419

14. Chromatographic retention–activity relationships for prediction of the toxicity pH-dependence of phenols / J. M. Bermúdez-Saldaña et al. *Chemosphere*. 2007. Vol. 69, no. 1. P. 108–117. URL: https://doi.org/10.1016/j.chemosphere.2007.04.060

15. Phenol. Medical Management Guidelines. Toxic Substance Portal | ATSDR. URL: https://wwwn.cdc.gov/TSP/MMG/MMGDetails.aspx?mmgid=144&toxid=27

16. Badria, Farid A. “Phenolic Compounds-Chemistry, Synthesis, Diversity, Non-Conventional Industrial, Pharmaceutical and Therapeutic Applications.” (2022).

17. Artificial neural networks: applications in chemical engineering / M. Pirdashti et al. *Reviews in Chemical Engineering*. 2013. Vol. 29, no. 4. URL: https://doi.org/10.1515/revce-2013-0013

18. Goncalves V., Maria K., da Silv A. B. F. Applications of Artificial Neural Networks in Chemical Problems. *Artificial Neural Networks - Architectures and Applications*. 2013. URL: https://doi.org/10.5772/51275.

19. Agatonovic-Kustrin S., Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*. 2000. Vol. 22, no. 5. P. 717–727. URL: https://doi.org/10.1016/s0731-7085(99)00272-1

20. Mutihac L., Mutihac R. Mining in chemometrics. *Analytica Chimica Acta*. 2008. Vol. 612, no. 1. P. 1-18. URL: https://doi.org/10.1016/j.aca.2008.02.025

21. Pushkarova Y., Kholin Y. The classification of solvents based on solvatochromic characteristics: the choice of optimal parameters for artificial neural networks. *Open Chemistry*. 2012. Vol. 10, no. 4. P. 1318−1327. URL: https://doi.org/10.2478/s11532-012-0060-z

22. Benyelloul K., Aourag H. Elastic constants of austenitic stainless steel: Investigation by the first-principles calculations and the artificial neural network approach. *Computational materials science*. 2013. Vol. 67. P. 353−358. URL: https://doi.org/10.1016/j.commatsci.2012.09.005

23. Identification of water samples from different springs and rivers of Kharkiv: Comparison of methods for multivariate data analysis / Y. N. Pushkarova et al. *Moscow University Chemistry Bulletin*. 2013. Vol. 68, no. 1. P. 60−66. URL: https://doi.org/10.3103/S0027131412060077

24. Pushkarova Y., Panchenko V., Kholin Y. Application an Artificial Neural Network for Prediction of Substances Solubility. *IEEE EUROCON 2021 − 19th International Conference on Smart Technologies*, Lviv, Ukraine, 6–8 July 2021. URL: https://doi.org/10.1109/EUROCON52738.2021.9535593

25. Mateo J., Rieta J. J. Radial basis function neural networks applied to efficient QRST cancellation in atrial fibrillation. *Computers in biology and medicine*. 2013. Vol. 43, no. 2. P. 154−163. URL: https://doi.org/10.1016/j.compbiomed.2012.11.007

26. Koç M. L., Özdemir Ü., İmren D. Prediction of the pH and the temperature-dependent swelling behavior of Ca^{2+}-alginate hydrogels by artificial neural networks. *Chemical engineering science*. 2008. Vol. 63, no. 11. P. 2913−2919. URL: https://doi.org/10.1016/j.ces.2008.03.012

27. Venkateswarlu C., Rao K. V. Dynamic recurrent radial basis function network model predictive control of unstable nonlinear processes. *Chemical Engineering Science*. 2005. Vol. 60, no. 23. P. 6718−6732. URL: https://doi.org/10.1016/j.ces.2005.03.070

28. Broomhead D. S., Lowe D. Multivariable functional interpolation and adaptive networks, complex systems, vol. 2. 1988.

29. Description and simulation of dynamic mobility networks. / A. Scherrer et al. *Computer Networks*. 2008. Vol. 52, no. 15. P. 2842−2858. URL: https://doi.org/10.1016/j.comnet.2008.06.007

30. Tylman W., Anders G. J. Application of probabilistic networks for decision support in power system analysis. *Energy*. 2006. Vol. 31, no. 14. P. 2874−2889. URL: https://doi.org/10.1016/j.energy.2005.10.027

31. Scarselli F., Tsoi A. C. Universal approximation using feedforward neural networks: A survey of some existing methods, and some new results. *Neural networks*. 1998. Vol. 11, no. 1. P. 15−37. URL: https://doi.org/10.1016/S0893-6080(97)00097-X

32. Zheng P., Harrington P. D. B., Davis D. M. Quantitative analysis of volatile organic compounds using ion mobility spectrometry and cascade correlation neural networks. *Chemometrics and Intelligent Laboratory Systems*. 1996. Vol. 33, no. 2. P. 121−132.

33. Marini F., Zupan J., Magrì A. L. Class-modeling using Kohonen artificial neural networks. *Analytica Chimica Acta*. 2005. Vol. 544, no. 1-2. P. 306−314. URL: https://doi.org/10.1016/j.aca.2004.12.026

34. Pushkarova Y., Kholin Y. A procedure for meaningful unsupervised clustering and its application for solvent classification. *Central European Journal*

of Chemistry. 2014. Vol. 12, no. 5. P. 594−603. URL: https://doi.org/10.2478/s11532-014-0514-6

35. Hu L., Chen G., Chau R. M. W. A neural networks-based drug discovery approach and its application for designing aldose reductase inhibitors. *J. Mol. Graph. Model.* 2006. Vol. 24. P. 244−253. URL: https://doi.org/10.1016/j.jmgm.2005.09.002

36. Ligand- based virtual screening procedure for the prediction and the identification of novel β-amyloid aggregation inhibitors using Kohonen maps and Counterpropagation Artificial Neural Networks / A. Afantitis et al. *Eur. J. Med. Chem.* 2011. Vol. 46. P. 497−508. URL: https://doi.org/10.1016/j.ejmech.2010.11.029

37. Towards improved penicillin fermentation via artificial neural networks / C. Di Massimo et al. *Comput. Chem. Eng.* 1992. Vol. 16. P. 283−291.

38. Palancar M. C., Aragón J. M., Torrecilla J. S. pH-Control System Based on Artificial Neural Networks. *Ind. Eng. Chem. Res.* 1998. Vol. 37. P. 2729−2740. URL: https://doi.org/10.1021/ie970718w

39. Takayama K., Fujikawa M., Nagai T. Artificial Neural Network as a Novel Method to Optimize Pharmaceutical Formulations. *Pharm. Res.* 1999. Vol. 16. P. 1−6. URL: https://doi.org/10.1023/a:1011986823850

40. Vracko M. Kohonen Artificial Neural Network and Counter Propagation Neural Network in Molecular Structure-Toxicity Studies. *Curr. Comput-Aid. Drug*. 2005. Vol. 1. P. 73−78. URL: https://doi.org/10.2174/1573409052952224

41. Nandi S., Vracko M., Bagchi M. C. Anticancer activity of selected phenolic compounds: QSAR studies using ridge regression and neural networks. *Chem. Biol. Drug Des.* 2007. Vol. 70. P. 424−436. URL: https://doi.org/10.1111/j.1747-0285.2007.00575.x

42. Computational neural network analysis of the affinity of lobeline and tetrabenazine analogs for the vesicular monoamine transporter-2 / F. Zheng et. al. *Bioorg. Med. Chem.* 2007. Vol. 15. P. 2975−2992. URL: https://doi.org/10.1016/j.bmc.2007.02.013

43. Schultz T.W. TETRATOX: Tetrahymena population growth impairment endpoint − a surrogate for fish lethality. *Toxicol. Meth*. 1997. Vol. 7. P. 289–309.

44. Beale M. H., Hagan M. T., Demuth H. B. Neural Network Toolbox. Getting started guide. Natick, MA: The MathWorks, Inc. 2018. 136 p.

45. Beale M. H., Hagan M. T., Demuth H. B. Neural Network Toolbox. User's guide. Natick, MA: The MathWorks, Inc. 2018. 558 p.

46. Demuth H., Beale M. Neural Network Toolbox for use with MATLAB. Natick, MA: The MathWorks, Inc. 2004. 846 p.

47. Nguyen D., Widrow B. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *IEEE 1990 IJCNN International Joint Conference on Neural Networks,* San Diego, CA, USA, 17−21 June 1990. URL: https://doi.org/10.1109/IJCNN.1990.137819

48. Balabin R. M., Safieva R. Z., Lomakina E. I. Universal technique for optimization of neural network training parameters: gasoline near infrared data example. *Neural Computing and Applications*. 2009. Vol.18, no. 6. P. 557−565.

49. Chetouani Y. Change detection in a distillation column based on the generalized likelihood ratio approach. *Journal of Chemical Engineering & Process Technology*. 2011. Vol. 2, no. 5. 1000115. URL: https://doi.org/10.4172/2157-7048.1000115

50. Shen W., Guo X., Wu Ch., Wu D. Forecasting stock indices using radial basis function neural networks oprimized by artificial fish swarm algorithm. *Knowledge-based Systems*. 2011. Vol. 24, no. 3. P. 378−385. URL: https://doi.org/10.1016/j.knosys.2010.11.001

SUPPLEMENTARY INFORMATION

CERTIFICATE

Google Scholar
Bowker

INTERNATIONAL SCIENCE GROUP

OUCI
Open Ukrainian Citation Index
doi
Crossref Content Registration

is awarded to

Mustafa Al Saker

for active participation
XII International Scientific and Practical Conference
"ACTUAL PRIORITIES OF MODERN SCIENCE, EDUCATION AND PRACTICE"
March 29 – April 01, 2022, Paris, France
24 Hours of Participation
(0,8 ECTS credits)

Organizing committee Ekaterina Zvereva

COLLECTION OF SCIENTIFIC PAPERS
SCIENTIA

CERTIFICATE OF PARTICIPATION

Mustafa Al Saker

participated in the III International Scientific and Theoretical Conference
THEORY AND PRACTICE
OF MODERN SCIENCE

April 1, 2022
Kraków, Republic of Poland

The conference is included in the Academic Resource Index ResearchBib catalog;

Scan the code to get access to the conference proceedings

EUROPEAN SCIENTIFIC PLATFORM

Head of the European Scientific Platform
Chairman of the Organizing committee
MARIIA HOLDENBLAT

Conference proceedings are publicly available under terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).